



Dokumentacija: Algoritam za preporuku hotela

1. Dohvat podataka

- Iz baze podataka dohvaćaju se svi **hoteli** i njihove **recenzije**.
- U obzir se uzimaju samo recenzije koje su **odobrene** i nisu **obrisane**.

2. Analiza trenutnog korisnika

- Identifikuju se svi hoteli koje je trenutni korisnik već ocijenio.
- Kreira se lista ID-jeva tih hotela, koja će kasnije služiti za filtriranje.

3. Izračun dinamičkog praga

- Dinamički prag određuje minimalnu vrijednost ocjene koja se uzima u obzir.
- Prag se prilagođava na osnovu prosječne ocjene trenutnog korisnika:
 - Ako korisnik prosječno ocjenjuje visoko (npr. **4.5**) → prag se postavlja na **4.0**.
 - Ako korisnik prosječno ocjenjuje nisko (npr. **2.5**) → prag se postavlja na minimalnih **3.0**.

4. Pronalaženje sličnih korisnika

- Slični korisnici se definišu kao oni koji su ocijenili barem jedan hotel koji je ocijenio i trenutni korisnik.
- Ograničenje: Koristi se **jednostavan pristup zasnovan na preklapanju hotela** (bez dodatnog ponderisanja).

5. Dohvat recenzija sličnih korisnika

- Iz baze se povlače sve recenzije koje su ostavili slični korisnici.
- Kao i ranije, u obzir se uzimaju samo **odobrene recenzije**.

6. Izračun vremenskih težina

- Svako recenziji se dodjeljuje težina na osnovu starosti.
- Formula:

$$\text{weight} = 2^{-\frac{\text{days}}{180}} \quad \text{weight} = 2^{-180 \text{days}}$$

- Primjeri:
 - Recenzija od jučer: ≈ 1.0
 - Recenzija od 6 mjeseci: **0.5**
 - Recenzija od 1 godine: ≈ 0.25

7. Izračun ponderiranih prosječnih ocjena

- U obzir se uzimaju samo hoteli koje korisnik **nije ocijenio**.
- Za svaki hotel računa se ponderisana prosječna ocjena:

$$\text{WeightedAvgRating} = \frac{\sum (\text{rating} \times \text{weight})}{\sum \text{weight}}$$

- Dodatno se bilježe:
 - Ukupan broj recenzija
 - Ukupna težina svih recenzija

8. Rangiranje i odabir preporuka

- Hoteli se rangiraju prema sljedećim prioritetima:
 1. **Ponderisana prosječna ocjena** (najvažnije)
 2. **Ukupna težina recenzija** (prednost novijim recenzijama)
 3. **Broj recenzija** (stabilnost rezultata)

9. Fallback mehanizam

- Aktivira se kada:
 - Nema dovoljno sličnih korisnika
 - Nema dostupnih preporuka
- U tom slučaju algoritam vraća **globalno najbolje hotele** po prosječnoj ocjeni, koristeći isti dinamički prag.

☒ Ovaj algoritam predstavlja **user-based collaborative filtering** pristup sa sljedećim specifičnostima:

- Dinamički prag zasnovan na profilu korisnika
- Vremensko ponderisanje recenzija (novije imaju veću vrijednost)
- Jednostavna definicija sličnosti korisnika (preklapanje ocijenjenih hotela)
- Fallback mehanizam za osiguranje preporuka čak i kada nema sličnih korisnika